1. Record Nr. UNINA9910481953403321

Autore	Bressoud Thomas
Titolo	Introduction to data systems : building from Python / / Thomas Bressoud, David White
Pubbl/distr/stampa	Cham, Switzerland : , : Springer, , [2020] Â©2020
ISBN	3-030-54371-4
Edizione	[1st ed. 2020.]
Descrizione fisica	1 online resource (XXIX, 828 p. 81 illus., 65 illus. in color.)
Disciplina	006.312
Soggetti	Data mining Python (Computer program language)
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Part I Foundation -- 1. Introduction -- 2. File Systems and File Processing -- 3. Python Native Data Structures -- 4. Regular Expressions -- Part II Data Systems: The Data Models -- 5. Data Systems Models -- 6. Tabular Model: Structure and Formats -- 7. Tabular Model: Access Operations and pandas -- 8. Tabular Model: Advanced Operations and pandas -- 9. Tabular Model: Transformations and Constraints -- 10. Relational Model: Structure and Architecture -- 11. Relational Operations: Single Table -- 12. Relational Operations: Multiple Tables -- 13. Relational Database Programming -- 14. Relational Model: Design, Constraints, and Creation -- 15. Hierarchical Model: Structure and Formats -- 16. Hierarchical Model: Operations and Programming -- 17. Hierarchical Model: Constraints -- Part III Data Systems: The Data Sources -- 18. Overview of Data Systems Sources -- 19. Networking and Client-Server -- 20. The HyperText Transfer Protocol -- 21. Interlude: Client Data Acquisition -- 22. Web Scraping -- 23. RESTful Application Programming Interfaces -- 24. Authentication and Authorization.
Sommario/riassunto	Encompassing a broad range of forms and sources of data, this textbook introduces data systems through a progressive presentation. Introduction to Data Systems covers data acquisition starting with local files, then progresses to data acquired from relational databases, from REST APIs and through web scraping. It teaches data forms/formats

from tidy data to relationally defined sets of tables to hierarchical structure like XML and JSON using data models to convey the structure, operations, and constraints of each data form. The starting point of the book is a foundation in Python programming found in introductory computer science classes or short courses on the language, and so does not require prerequisites of data structures, algorithms, or other courses. This makes the material accessible to students early in their educational career and equips them with understanding and skills that can be applied in computer science, data science/data analytics, and information technology programs as well as for internships and research experiences. This book is accessible to a wide variety of students. By drawing together content normally spread across upper level computer science courses, it offers a single source providing the essentials for data science practitioners. In our increasingly data-centric world, students from all domains will benefit from the “data-aptitude” built by the material in this book.